DOCUMENT RESUME

ED 186 433 SP 016 111

AUTHOR Porter, Nancy M.: Eileenchild, Margaret T.
TITLE The Effectiveness of Women's Studies Teaching.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Feb 80
CONTRACT NIE-P-79-0017
NOTE 81p.

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Curriculum Development: *Educational Needs; Educational Philosophy: Feminism: *Humanistic Education: *Needs Assessment: Outcomes of Education; Participant Satisfaction: *Program Effectiveness; *Progressive Education: Sociocultural Patterns; Teacher Effectiveness: *Womens Studies

ABSTRACT

This planning study reviews the literature on teaching effectiveness in women's studies in the context of several integrative studies in research on teaching and research on the evaluation of teaching. The study suggests that changes in the cognitive development of women's studies students are as fully important as changes in sex-role attitudes. An approach to research and evaluation which acknowledges the interaction between students, teachers, and the subject matter, and which seeks to establish relationships between multiple variables involved in an interdisciplinary area such as women's studies, is recommended. (Authors/LH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

THE EFFECTIVENESS OF WOMEN'S STUDIES TEACHING

ED186433

Nancy M. Porter
Portland State University

Margaret T. Eileenchild
Harvard Graduate School of Education

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

U.S. Department of Health, Education, and Welfare
Patricia Roberts Harris, Secretary

National Institute of Education
P. Michael Timpane, Acting Director
Program on Teaching and Learning
Lois-ellin Datta, Associate Director

February 1980

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402

SP 016 111

JUN 9 1980

This study was conducted for the National Institute of Education, Department of Health, Education, and Welfare, under purchase order number NIE-P-79-0017. The opinions expressed in this report do not necessarily reflect NIE position or policy, and no official endorsement by NIE or HEW should be inferred.

PREFACE

This monograph is part of a series of eight, commissioned by the National Institute of Education (NIE), following recommendations presented in Seven Years Later: Women's Studies Programs in 1976 by Florence Howe.

Seven Years Later was commissioned and published by the National Advisory Council on Women's Educational Programs. It was the first federally financed investigation of women's studies. Howe visited and reported on 15 mature women's studies programs, in the context of national trends and directions within women's studies. She focused on some areas (funding and faculty, for example, in 4-year colleges and universities) and excluded others (community colleges, for example) entirely.

The eight new monographs take up some of the questions Howe explored; they also initiate explorations of untouched areas. All are concerned primarily with higher education.

Each monograph reviews literature relevant to its focus produced during this first decade of women's studies. Each also attempts to provide definition to a field in the process of development, a complex interdisciplinary area of scholarship and curriculum that is also a strategy for educational change. Beyond definition, each monograph recommends research essential for the future development of women's studies and important to educational research more broadly.

While producing these eight monographs, the researchers met for 2 days to compare perspectives and conclusions. As a group, we agreed to offer four formal recommendations to the National Institute of Education and to all other Federal agencies and private foundations that might pursue and support research and development in women's studies.

1. The need for future monographs. We recognize the importance of continuing the development of information about relatively unexplored aspects of women's studies. Among those aspects that need examination, the following are especially important:

   a. Graduate programs and graduate education in women's studies.

   b. Schools of education and women's studies activities.

   c. Women's studies in the elementary and secondary schools.

   d. Noncredit curriculum development in women's centers, continuing education for women, and community-based centers.

2. The need for a data base on women's studies. All eight monographs recommend the development of a data base on women's studies as a necessary prelude to research design and activity. Since women's studies has developed in different ways and at different rates on almost all campuses throughout the country, such a data base must be longitudinal in design. While we are recommending its development in higher education, such a data base eventually will need to be extended to all areas of education. We recommend that annual updates be built into the data base to ensure its proper maintenance and continued usefulness.

3. The qualifications of research personnel in women's studies. We believe that the distinctive and complex nature of women's studies, as well as the fact that it is still a developing field, makes it essential that researchers be thoroughly informed about

women's studies and be prepared to involve women's studies practitioners in every aspect of research design and process. We are hopeful that NIE and other Federal agencies, private foundations, and research institutions will place responsibility for research in women's studies in the hands of those experienced and knowledgeable about the area.

4. Considerations significant to evaluation research and methodology. We believe that the evaluation of women's studies in higher education will be an important activity of the coming decade. Such evaluation will serve the researcher interested in the processes of changing higher education, as well as those charged with administering institutions. Such evaluation should, on principle, also clearly serve the developmental needs of women's studies programs, engaging them in the design and processes of research, as well as in the determination of the uses to which research will be put. Evaluations should be conducted onsite, with the cooperation of program participants, and within a framework of longitudinal data about that program and others nationally. Researchers and evaluators should be sensitive to the work of the National Women's Studies Association in this regard, and its activities serving the needs of women's studies programs.

Detailed, further recommendations appear in each of the eight monographs.

We wish to express our appreciation to the National Institute of Education for pursuing this

research, and particularly to C.B. Crump, who directed the Women's Studies Planning Studies project.

Christine Bose
Kathryn Brooks
Margaret Eileenchild
Allana Elovson
Mareyjoyce Green
Blanche Glassman Hersh
Florence Howe
Paul Lauter
Elizabeth Ness Nelson
Nancy Porter
Janet Priest-Jones
Elaine Reuben
Mary Jo Strauss

SUMMARY

This planning study reviews the literature on teaching effectiveness in women's studies in the context of several integrative studies in research on teaching and research on the evaluation of teaching.

To date, the experimental research in women's studies has, for the most part, concentrated on assessing changes in attitude toward sex roles and sex stereotyping as the appropriate objectives for measuring teaching effectiveness in women's studies. We suggest that cognitive development, an area that has not been addressed in the literature, is at least as important as attitude change. An approach to research and evaluation that acknowledges the interaction between students, teachers, and the subject matter and seeks to establish relationships between multiple variables that obtain in teaching and learning is recommended.

We propose both long- and short-range goals in research. These include research that will contribute to a knowledge base and acquaint the larger educational community with women's studies approaches to teaching and learning, as well as formative studies that will be of direct use to women's studies practitioners.

Nancy M. Porter
Portland State University

Margaret T. Eileenchild
Harvard Graduate School of Education

CONTENTS

	Page
PREFACE	iii
SUMMARY	vii
1. INTRODUCTION	1
2. RESEARCH ON TEACHING	5
Overview	5
The State of the Art	6
3. EVALUATION OF TEACHING	13
Overview	13
The State of the Art	15
Summary	19
4. REVIEW OF THE LITERATURE ON THE EVALUATION OF TEACHING EFFECTIVENESS IN WOMEN'S STUDIES	21
Overview of Approaches	21
Review of the Literature	26
Women's Studies Values and Teaching: Theory and Practice in the Classroom	47
Summary	52
5. CONCLUSIONS AND RECOMMENDATIONS	55
Research on Teaching	55
Evaluation and Improvement of Teaching	60
BIBLIOGRAPHY	65
APPENDIX	75

1. INTRODUCTION

Thinking about the questions of effective teaching in women's studies has led us to consider two more basic questions: What is teaching? What is learning? Both questions inevitably lead back to: What is education?

Our philosophies of education, as women's studies teachers, students, and administrators, influence our notions of what we might want to teach or study, why we might want to teach or study one thing over another, how we might go about the process of teaching and learning, and to what ends our instruction and study might be put. Teaching and learning, whether effective or ineffective, are the interplays between the subject matter (the whats and whys), the approaches (the hows), and the purposes (the ends).

Questions of teaching effectiveness, in or out of women's studies, are embedded in the various philosophies of education that influence and socialize both teaching and learning. The objectives of teaching, learning, and evaluation research ultimately trace back to the articulated and unarticulated philosophies from which the various objectives arise. And the methodologies for assessing effective teaching are linked to the varying and sometimes conflicting objectives held by students, teachers, and researchers.

The literature that surrounds the development of women's studies courses and programs often emphasizes the unique promise of the field. Given this, it should be realized that women's studies shares in at least three traditions of higher education in the United States: progressive or relevant education; the fostering of scholarship; and the preparation of students for careers and professions.

By relevant or progressive, we mean education that is viewed as the "fundamental method of social

progress and reform," a process that integrates the psychological and the sociological in cognition and affect, and a curriculum that moves in the "positive direction of providing a body of subject matter" that is "richer, more varied and flexible" than the traditional when "judged in terms of the experience of those being educated" (John Dewey, 1897).

By the fostering of scholarship, we mean the interpretation, criticism, and communication of the accumulated knowledge of a number of disciplines, as well as the generation of new knowledge within and across the disciplines. And by the preparation of students for work and the professions, we mean the particular emphasis given to the creation of opportunities within the general curriculum for students to explore possible fields of work, as well as the opportunity to acquire specific knowledge and skills useful in career and professional development.

Women's studies, as an emerging discipline, overlays these traditional concerns with an emphasis on feminism and, in theory at least, an interdisciplinary perspective in the interpretation and criticism of the accumulated knowledge of a number of academic disciplines and within women's studies itself. Women's studies is also a political phenomenon in that it is viewed by many of its practitioners as an outgrowth of various social change movements of the 1960's, particularly the women's liberation movement.

Feminism, as we use the term in the context of the classroom, is the desire to increase the power and autonomy of women as individuals and as a group through instructional processes that enable them to acquire knowledge and skills to make informed, flexible choices in their educations and their lives. Feminism, in this sense, may be viewed either for its own sake or as a way station on the road to a more responsible and informed humanism.

In an extension of the definition of "interdisciplinary" offered in Signs (1975), we use this term to mean an approach to subject matter and course design in which one person skilled in several disciplines pursues one subject, or several teachers each skilled in a single discipline explore a single subject together (Yates, 1977), or students and teachers together use methods and information from the research or scholarship of multiple disciplines. And by political, we mean not only the field of women's studies viewed as the academic arm of the women's movement, but the process by which women's studies teachers and students strive to make their classes a mediation between what has been and is and what is and what might be in education, culture, and society.

It is this overlay of feminism, the interdisciplinary, and the political that makes the assessment of teaching and learning, and the structures in which they take place, relatively problematic for the participant and the observer who have been trained in more conventional modes.

Because the definition and assessment of teaching effectiveness in women's studies is complex, it has seemed worthwhile to us to introduce the published, unpublished, and informal literature of women's studies with discussions of the status of research on teaching and on the evaluation of teaching. Our recommendations for further research in the evaluation of teaching and learning in women's studies are drawn from this comprehensive body of research.

Finally, in an effort to establish a more comprehensive statement of the values of women's studies teaching -- that is, the goals of women's studies teaching that inform specific course objectives -- we have surveyed women's studies teachers across the country. The results of this survey are included in an appendix to this monograph.

2. RESEARCH ON TEACHING

Overview

Research on teaching has a respectable longevity, at least in the United States. The earliest research, reported by Doyle (1975), was an 1896 study concerned with the factors that contribute to effective teaching.

Dubin and Traveggia (1968) report from a 1924 study, entitled "The Lecture vs. the Class Discussion Method of College Teaching," that differences in method produce no significant differences in student learning as measured by final examination scores. Dubin and Traveggia reviewed 90 additional studies and concluded that no one teaching method is superior to another when student learning is measured by examination scores. They point out, however, that these scores do not reflect all of the learning that takes place during a course and that educational goals such as personal and intellectual development are also important.

Dubin and Traveggia's unequivocal conclusion notwithstanding, the volume of research in the area of effective teaching, and particularly research related to effective teaching at the college and university level, is enormous and growing. This growth may be attributed to the cumulative effects of an expanding knowledge and data base, to the movement for accountability in higher education, and to this movement's offspring, faculty development, which has the improvement of teaching as a major component.

The literature on effective teaching is so vast (Gage, 1972) that even a cursory review is beyond the scope of this planning study. However, we believe that it is useful to review below four integrative papers that are concerned with conceptual rather than methodological problems in research on teaching. Such

seems an appropriate strategy since the state of the literature on the effectiveness of women's studies teaching is at the conceptual stage, as will be seen in section 4.

The State of the Art

As Biddle (1964), Smith (1971), and Doyle (1975), respectively, point out, there are problems with the "facts" of teaching effectiveness, the "concepts," and the relationship between teaching and learning.

Biddle comments that, in the "literally thousands of studies...few if any 'facts' seem to have been established concerning teaching effectiveness, no method of measuring competence has been approved, and no methods of promoting teacher adequacy have been widely adopted." With respect to the concepts of teaching effectiveness, Smith notes: "Despite all of our efforts, we apparently have no generally accepted conceptual system, psychological or otherwise, by which either to formulate or to identify skills of teaching."

And Doyle concludes that the relationship between teaching and learning eludes the researcher. Efforts to achieve a "cumulative integration of research findings...have, with remarkable regularity, failed to support the existence of stable and consistent relationships between teacher variables and effectiveness criteria." Although Doyle finds some cause for "cautious optimism" in the findings of Flanders and Simon (1969) and Gage (1972), he warns that "past experience in this area strongly suggests that establishing stable teacher effectiveness relationships will not be easy."

Biddle: Variables in Effective Teaching

Biddle (1964) begins his review of research on the effectiveness of teaching with a conceptual

discussion of the basic variables involved in effective teaching. Defining teacher effectiveness as an "individual teacher's ability to produce agreed upon results," the author offers a model that includes seven variables that should be considered in research on teaching effectiveness: (1) the formative experiences of the teacher -- educational background, socialization, sex, age, or race; (2) the teacher "properties," that is, the specific skills, knowledge, or motives that the teacher brings to teaching and learning; (3) the teacher "behaviors," that is, teaching style, responsiveness to students, mannerisms, and so on; (4) the immediate effects of formative experiences, properties, and behaviors on students; (5) the long-term effects of the first four variables on students; (6) the classroom environment itself, both the make-up of the class and the physical environment; and (7) the institutional or community context in which teaching and learning take place.

Biddle suggests possible methods for measuring the seven variables, including observational techniques, student ratings, achievement tests, and the self-reports of students. He calls for the integration of the research on teaching, stating that the measurement of variables has become confused with the variables themselves. He contends that all of the variables must be considered if the confusion is to be reduced and if we are to understand the complexity of the problem.

Biddle's concerns that there must be agreement about objectives, that both the short- and long-term effects of instruction on students, as well as the institutional and/or community context, must be taken into account, and that variables must be integrated are particularly useful in conceptualizing research on the effectiveness of teaching in women's studies. However, he fails to include the learner as an important variable in the process of measurement.

Henderson and Lanier: Value of Teacher's Goals

Henderson and Lanier (1972) focus on the lack of conceptualization of effective teaching and the consequent absence of systemization in the research. Their work is unique in that they propose assessing the value of the teacher's goals or objectives, that is, the purpose of the content of the instruction.

Henderson and Lanier discuss the social responsibility of teachers as an important variable in assessing effectiveness. However extreme, the following example illustrates the point. If the end of a course in atomic physics were to enable individuals to build an atomic bomb, then regardless of the means employed, no matter how well the course was taught, the result would not be effective teaching. As we know, many sexist and misogynist courses are very well taught.

Doyle: "Process-Product" Paradigm

Doyle (1975) uses Kuhn's (1970) concept of paradigm (a framework that includes the questions that are asked in a particular field, the methods that are used to ask the questions, and the tools that are used to evaluate the answers) to define the approach most widely used in the research on teaching effectiveness: the "process-product" paradigm. Doyle argues convincingly that questions of teaching effectiveness must be reformulated within alternative paradigms if researchers are to do more than continue to produce negative or nonsignificant results. The production of these negative or nonsignificant results dominates the research on teaching effectiveness in general and, as discussed in section 4, in women's studies in particular.

Operationally, the process-product paradigm focuses on isolating statistical associations between the frequency of specific teacher behaviors (the independent variable) and changes in class mean scores (the dependent variable). Although a good deal of the

research is of the descriptive, correlational variety, there is an assumption within the paradigm that the behaviors of teachers have a causal impact on the behavior of students. Not only is this assumption of causality not warranted from the data available, Doyle points out that some of the research actually indicates that the behavior of students is the cause of the behavior of teachers.

Doyle suggests two alternatives to the process-product paradigm: the student mediating process (or process-process) paradigm, and the "culture of the school" paradigm. Both have important implications for the assessment of teaching effectiveness in women's studies.

In the student mediating paradigm, variations in the outcomes of student learning are viewed as a function of the mediating processes employed by the students during the learning process itself. The mediating process is influenced in part, but not exclusively, by instruction. This paradigm moves the student from the passive, acted-upon role assumed in the process-product paradigm into the arena as an interactor. Such variables as the ability and cognitive styles of the students mediate between the teacher and the outcomes of the learning.*

*A note on terminology is called for. The terms learning style and cognitive style are sometimes used interchangeably in the literature. As we use the term, however, learning style is to be interpreted as defining the behaviors of students that are related primarily to personality and that can be observed or inferred from behavior. For instance, a student may be a cooperative, a compliant, a highly-motivated, or a competitive learner. Cognitive style refers to those aspects of thinking that are, for the most part, not observable but are preferred styles of information-seeking and information-processing in solving problems (Coop and Sigel, 1971). Freize et al. (1978) offer a current discussion of sex differences in cognitive style.

Not only does the process-process paradigm reflect more accurately the reality of the classroom, it could be expanded to include the student learning that takes place outside of the classroom as well. We need to be reminded that the classroom is only one component of a course: it provides a context for the reading, discussion, thinking, and writing that go on outside (Stizel, 1977). Further, in women's studies, students sometimes engage in field work, and some actively participate in program maintenance, including service on curriculum committees and teaching. These activities need to be considered in understanding the complex influences on the classroom.

The bidirectional approach of the process-process paradigm is important to research on teaching effectiveness in women's studies, particularly as it relates to the behavior of students on such variables as competence, motivation, and enthusiasm, which, in turn, influence the teacher and affect the process of teaching and learning. And as an additional consideration, because the majority of women's studies teachers and students are female, current research on women as learners, both in terms of learning and cognitive styles, should be taken into account in further research in this area.

Doyle's second alternative hypothesis, the culture of the school paradigm, includes classroom variables not connected specifically to the behavior of teachers or students. Rather than isolating variables, this paradigm focuses on the use of the inductive-observational method to record and analyze the full range of variables influencing a particular class. Often used for formulating hypotheses, this method seems an appropriate methodology for women's studies at this time. As noted in section 4, the descriptive reviews, case studies, and personal narratives of women's studies teachers and students are useful in conceptualizing the variables of effective teaching in women's studies.

Winne and Marx: Reconceptualization of Research

Winne and Marx (1977) build on Doyle's process-process paradigm and further support Doyle's analysis that there are serious conceptual and methodological problems with descriptive correlational and experimental research that looks for causal relationships between the behavior of teachers and the learning of students. They propose a reconceptualization of research on teaching to include variables that describe "how one learns." They further propose that these variables should apply to students and teachers alike. In their view, "an adequate knowledge about teacher effectiveness cannot develop without considering the mental life of teachers and students."

Winne and Marx see the teacher as primarily, although not exclusively, responsible for communicating. Such communication involves choices about the structure of the subject matter, the cognitive and learning styles of the teacher, and "dynamic decisionmaking" on the part of the teacher in the classroom. Students are primarily, but not exclusively, responsible for acquiring skills and knowledge, and this also involves the cognitive and learning styles and the dynamic decisionmaking of the students.

Although the authors call for research that takes into account the multivariate and bidirectional nature of the teaching and learning process, they acknowledge that making these concepts operational is difficult. Cognitive processes can only be inferred from behavior, and the full extent and actual structure of what one learns or how one thinks -- whether as teacher or as student -- is nearly impossible to quantify. Nevertheless, the approach of Winne and Marx is useful for providing a comprehensive framework for the analyses, case studies, and descriptive research in women's studies. Their model suggests that students and teachers share responsibility for the quality of the teaching and learning process.

As a final note, Winne and Marx view the choices of subject matter that teachers make within a framework that preserves the "structure of the subject" (Bruner, 1963). In women's studies, the body of knowledge is developing rapidly, and teachers and students are faced with a knowledge explosion. The body of knowledge is also becoming more organized, and the curricular materials more sophisticated. The decisions that teachers make, therefore, about the design and organization of a course or a curriculum are also dependent upon the body of knowledge that is available. And this in turn influences teaching effectiveness.

3. EVALUATION OF TEACHING

Overview

Research paradigms, at least of the conceptual rather than the methodological variety, favor a multivariate and bidirectional approach to understanding and assessing effective teaching and learning. Conceptual papers on the evaluation of teaching also take a comprehensive approach. The position paper of the American Association of University Professors (AAUP, 1974) calls for the use of multiple measures to assess the effectiveness of instruction. Measures of student learning, student ratings of instruction, student and alumni interviews, classroom visitation by peers, instructor self-evaluation, and outside expert evaluations all are recommended.

However, most colleges and universities rely solely on student ratings for the evaluation of teaching. Although the research on teaching suggests that looking for correlational or causal relationships between the behavior of teachers and the achievement of students has not been particularly fruitful, the wide use of student ratings necessitates a review of current attitudes toward the reliability, validity, and usefulness of these ratings.

Terminologies, assumptions, and some confusion inherent in the studies reported on in this section need to be articulated. Reliability, for the most part, refers both to the extent to which an instrument is consistent in measuring objectives and to the consistency of the student raters. Validity, in most instances, does not refer to the validity of the instrument used. Instruments in these studies have been factor analyzed and validated. Validity, then, refers to the ability of the students to assess validly the objectives and properties that are assumed to characterize effective teaching. However, because it is often difficult to agree upon what is to be

measured and because different courses may have different objectives (and adaptations of instructional format to meet these objectives), questions of validity are often problematic. As Talburtt et al. (1977) comment, students seem less adept at assessing the outcomes of innovative as opposed to conventional courses or educational experiences.

We also need to be aware that all of the work reported on is correlational and indicates relationships between variables. As the adage goes, however, correlation does not imply causation. Similarly, the research assumes the existence of universal characteristics of effective teaching that cross and encompass the various academic disciplines, so that results from an economics class can be compared with results from a class in French language or one in educational psychology. Although Feldman and Newcomb (1969) indicate that students and teachers associated with different academic fields tend to have different values and characteristics of personality, Socklaff and Papacostas (1975) state that the discipline does not influence qualities of effective teaching. For the most part, researchers in the area assume the latter position.

Usefulness of student ratings is a mixed concept. Student ratings are sometimes used for either formative or summative evaluation.* Centra (1972) reports, however, that formative ratings do little, if anything, to improve teaching in the short run. We have found no studies describing the long-term effect of student ratings on the improvement of instruction. Centra hypothesizes that student ratings tend to tell the teacher what she or he already knows.

*Formative evaluation is intended to modify or improve instruction during a course. Summative evaluation is conducted at the end of the course, with the purpose of improving teaching in the future.

Often, student ratings are used for a combination of purposes in addition to the improvement of teaching. These include making personnal or administrative decisions and providing information for students in choosing classes and instructors. In some instances, student ratings merely take up space in inactive files.

Confusion clouds the literature when it becomes unclear whether the student rating forms measure the effectiveness of the teacher, the effectiveness of the classroom experience, or the effectiveness of the course as a whole. For example, if a student indicates "increased interest in the subject," the studies do not indicate whether this interest is related to the enthusiasm or skill of the teacher, the composition of the class, the background of the other class members, the quality of the reading materials, or the design of laboratory experiments.

The State of the Art

A thorough review of the evaluative research in teaching effectiveness is beyond the scope of this planning study. The following review focuses on the major questions of reliability, validity, and usefulness in student ratings.

Validity, Reliability, Usefulness Of Student Ratings

Rodin and Rodin (1972), in a frequently quoted study, assert that students "rate most highly instructors from whom they learn the least." The authors define the objective criterion as what students learn and the subjective criterion as how the students evaluate the instructor. If student ratings measure the effectiveness of teaching, there should be a positive correlation between objective and subjective criteria.

Rodin and Rodin determined that, with the initial ability of the students held constant, there was a negative correlation (-.746 $p < .05$) between what students learned as measured by problem-solving tests that exhausted the content of the course and their rating of the instructors. The authors argue that student evaluations reflect the personal and social qualities of the teacher rather than the instruction. "If how much students learn is considered to be the major component of good teaching, it must be concluded that good teaching is not validly measured by student evaluations in their current form."

In a critique of this research, Marsh et al. (1975, point out that there may have been methodological flaws. In a replication of the Rodin and Rodin study, with methodological problems compensated for and a greatly expanded list of variables, Marsh et al. found that student evaluations are both valid and useful.

The work of Costin et al. (1971), Frey et al. (1975), and McKeachie et al. (1971) supports the more moderate view of Marsh et al. over that of Rodin and Rodin.

Costin et al. report that student ratings are "reasonably" reliable when correlations from one year to the next and from mid- to end-of-semester are considered. They list a number of criteria that are often associated with effective teaching and learning by both students and teachers: thorough knowledge of the subject; preparation for class; motivation of students to do their best; presentations of new points of view or appreciations; and flexibility, enthusiasm, warmth, agreeableness, and friendliness. They conclude, however, that although a review of the empirical studies indicates that student ratings provide reliable and valid information, these ratings "fall far short of a complete assessment of an instructor's teaching contribution."

Frey et al. obtained student ratings from large multisection courses at three midwestern universities. Their results demonstrate that student accomplishment, clarity of teacher presentation, and organization or planning of the course correlate positively with final examination scores and are, therefore, valid indicators of effective teaching. In analyzing the relationship between student ratings and student characteristics, the authors found that the more experienced students tend to give higher ratings to their instructors.

McKeachie et al. gathered and analyzed the factors in a variety of student rating forms. Skill of the teacher, difficulty of the course, structure of the course, feedback to students, group interaction in the class, and rapport between students and teachers (warmth) emerged as six stable factors.

The authors conducted five empirical studies, involving more than 1,700 students, to determine the relationship between the above factors and effective teaching. With the caveat that none of the courses were identified as women's studies courses, certain of their findings need to be considered in the assessment of teaching effectivenss in women's studies.

In all five studies, teachers who were rated highly on "structure" tended to be more effective with women students than with men students. Teachers who were given high ratings on warmth or rapport tended to be effective on measures of student thinking (application) as opposed to measures of student knowledge (accumulation). And teachers whom students rated as having an impact on their beliefs were effective in changing attitudes. In four of the five studies, teachers rated highly on "skill" tended to be particularly effective with women students. McKeachie et al. recommend that results for men and women be analyzed separately: "what works for men does not necessarily work for women."

Influence of Sex Differences

Identification of sex differences in student ratings by McKeachie et al. also raises the question of the influence on student ratings of the sex of both the instructor and the student rater. Several studies that we have located are concerned with the issues of sex of the student, sex of the teacher, and sex stereotyping.

Ferber and Huber (1976) asked students to rate their college teachers in retrospect, and found that all women students evaluated all of their teachers more favorably than did men students. Harris (1975) found that female students generally rated their teachers' performance higher than did male students on descriptions of the style of the teacher. Defining a masculine teaching style as active, aggressive, and directive and a feminine teaching style as positive, facilitative, and listening, Harris found that teachers who used the masculine mode, irrespective of sex, were rated higher in performance than either male or female teachers who used the feminine mode.

In a subsequent experiment, Harris (1976) sought to test the applicability of research indicating that men and women tend to perceive male-stereotyped behavior as superior to and healthier than female-stereotyped behavior and that American women and men both tend to rate the performance and competence of women lower than that of men. Harris found that there was an "overwhelming" tendency for a teacher described in masculinely stereotyped terms to be rated more positively on all variables except warmth. The results of Harris's experiment suggest that feminine traits, rather than female gender, are viewed negatively in teaching.

Kashak (1978) determined that disciplines or subject areas traditionally associated with one or the other sex did not affect the ratings of male or female professors by students of either sex. The sex of the professor, however, seemed to be the crucial factor on

which males evaluated faculty members. Women students rated female professors equally with male professors on such variables as effective-ineffective, concerned-unconcerned, likeable-not likeable, and excellent-poor. They did not find female professors less powerful and did not discriminate on the basis of sex except to choose to take a course from a female teacher. Male students, on the other hand, assigned higher ratings to male professors on these variables, regardless of subject area. These results do not support the conclusion of Harris (1976) that it is stereotyped behavior, rather than sex itself, that is devalued.

Summary

The AAUP (1974) call for multiple measures for the assessment of teaching effectiveness has been heard throughout the land. And Bergquist and Phillips (1975) report that teacher self-evaluation and student learning style forms have been developed and are in use. Nevertheless, we have not been able to locate validation studies of these forms or of outside "expert" opinion.

Centra (1975) found that, in comparing student and instructor peer ratings of the same instructor, student raters, who have the reputation of being lenient (Hildebrand, 1972), are less lenient than faculty peers: the average colleague rating was 4.47 on a 5-point scale; the average student rating was 3.98. Centra concluded that peer ratings are not as valid as student ratings.

Our position with respect to the reliability, validity, and usefulness of student ratings is decidedly middle-of-the-road. The bulk of the literature supports the findings of Marsh et al. over those of Rodin and Rodin. It is clear, however, that factors such as the sex of the student and the sex of the teacher, and the issue of sexism or sex stereotyping, are potent variables that affect the use and

usefulness of student ratings. We agree with the AAUP that multiple measures should be accepted in the evaluation of teaching effectiveness. However, student ratings are better than nothing and are going to be around for awhile, and to eliminate them entirely because of their shortcomings would be to eliminate a large data base in and out of women's studies.

4. REVIEW OF THE LITERATURE ON THE EVALUATION OF TEACHING EFFECTIVENESS IN WOMEN'S STUDIES

This section covers three principal areas: (1) overview of the approaches taken in assessing teaching effectiveness in women's studies, who the investigators are, what methods they use, and what use is made of their findings; (2) chronological review of the published and unpublished literature during the 1972-1978 period; and (3) the published and informal literature of women's studies values, teaching methods, and classrooms.

Overview of Approaches

At least three different approaches have been used in assessing teaching and/or course or program effectiveness in women's studies, and a fourth is in the developmental stage. Although the practitioners of any one method have, to varying degrees, tried to incorporate some of the values, objectives, and methodologies of the other approaches, their work is based on fundamentally different premises. Overlapping values and objectives, and at times imprecise or conflicting terminologies, tend to blur distinctions in concept.

We categorize the four approaches as: (1) descriptive reviews of women's studies courses and programs, which provide context for asking more specific questions of teaching effectiveness; (2) student course ratings; (3) research of experimental design; and (4) exploratory or case study research.

Descriptive Reviews

In the early phases of course or program development, the descriptive review looks to document the need and value of women's studies. Factors such as who the students are, the relation between the content

of the courses and larger educational and societal issues, modes of instruction, and outcomes reported by students are assessed with an eye to establishing the course or program as academically respectable and socially necessary.

Examples of the descriptive review are numerous. They are often contained in program reports and requests for course credit or degree-granting status, and they are by no means limited to early women's studies. Wilson's review of the Women's Studies Program at the University of Puget Sound (1976), the final evaluation of "Images of Women in Literature" at the University of Minnesota (Davis et al., 1975-76), and Howe's Seven Years Later: Women's Studies Programs in 1976 (1977) are examples of this mode.

The people who do descriptive reviews may or may not also use the methods of inquiry traditional in social science research and may or may not test the extent to which courses and programs meet their objectives. As often as not, practitioners possess women's studies experience as teachers or administrators, or are sympathetic to women's studies. The descriptive review may contain elements of subjectivity. Certainly, to some extent, it must rely on impressions. It has been invaluable in establishing the institutional context in which women's studies teaching and evaluation occur.

A variation of the descriptive review -- the mandated evaluation by outside "experts" -- also occurs in women's studies. A comprehensive summary of either voluntary or mandated reviews is beyond the scope of this planning study, and many such documents are not generally available. The specific examples of the voluntary descriptive review that we include are those available to us that also measure the relative achievement of course or program objectives in terms of student course ratings or experimental design research.

Student Course Ratings

The second approach -- sometimes included in reports of the first -- adheres with more or less fidelity to the student course rating method of soliciting information on a standard departmental or all-university questionnaire, or on a form designed by an individual instructor or a women's studies program. The goals of this approach are to produce quantifiable data that can be used by decisionmaking groups: students, teachers, administrators, and personnel; and planning and curriculum committees.

Contrary to the findings of Yamoor et al. (1973), that formal course evaluation seemed not to have kept up with the growth of women's studies, 94 percent of the respondents to our survey reported that their courses were regularly evaluated (see appendix). Despite the stated reservations about the reliability, validity, and usefulness of student ratings, these ratings provide an unanalyzed data base for future research in teaching effectiveness. The research of the Women's Studies Evaluation Group at the University of Washington and the Measurement Services Center at the University of Minnesota are examples of extensive evaluation of course and program effectiveness.

Experimental Design Research

In the experimental design approach, a researcher or research team from the outside -- or from some point between outside and in -- administers tests to determine to what extent certain defined objectives have been achieved by the "treated" group. This approach may be used in descriptive reviews (Shueman and Sedlacek, 1976) or in case studies that also use student course evaluations (Davis et al., 1975-76), or it may be included in program self-survey reports (University of Minnesota, 1978).

Sometimes the research has hypotheses to prove (Borod, 1975) or seeks to validate the construct of

some particular instrument or scoring system (Coffman, 1978). In the most liberal vein, the researcher may function as a participant observer and/or check or supplement instrument data with interviews with students and teachers and -- the hardiest folk -- with content analysis of student course productions (Speizer, 1975). In studies based on the Multi-Attribute Utility Decision-Theoretic (MAUT) model, participants contribute to setting the goals to be measured (Brush et al., 1978).

Experimental design studies have, for the most part, sought to measure outcomes in relation to assumed or stated course or program objectives, used a pre- and post-test format (with or without a control group), gathered basic demographic information, and concentrated on measuring self-concept, sex stereotyping, and sex-role attitudes. Exceptions to and variation in this pattern will be noted.

It appears that experimental design studies either posit a connection between affective changes and the enhanced ability to absorb and utilize subject matter, or assume that affective change is one of the goals of the women's studies course. The use to which these studies are put is not always clear in the literature. Sometimes, the proposed course of action is to improve the design or scope of future studies. Sometimes, it is in the application of the results to specific issues in developing women's studies curriculum.

Exploratory or Case Study Research

The exploratory or case study approach differs from many of the experimental design research methods by more directly addressing the issues of materials, methodologies, structures, and pedagogies used in women's studies classrooms. Bonneparth (1978) describes the concept behind the development of this approach as follows:

> A commitment to innovation in curricula and teaching technique must be carried over to the evaluation process itself. Surveys of student characteristics and attitudes provide important feedback on many questions, but alternative evaluation techniques are needed to get at the theoretical issues...

Some of the theoretical issues raised by Bonneparth that are addressed in the literature on the exploratory or case study research include: the rationale behind establishing one kind of women's studies course or program rather than another; questions of content, focus, and progression; the relation between learning and doing; and the process of making goals operational. Again, it should be noted that there may be overlapping objectives between this approach and the others.

The research conducted by Talburtt et al. (1977), based on the formative evaluation model, focused on the effects of three types of women's educational programs (internships, women's studies courses, and skills development classes) in eight institutions. The exploratory research of Elovson and Cockcroft (1977) concentrated on assessing the impact of women's studies courses on students' lives. The case studies of Davis et al. (1975-76), Kritek and Glass (1978), and Register et al. (1978) assessed the curricular, methodological, structural, pedagogical, and outcome aspects of, respectively, an "Images of Women in Literature" course, a course oriented toward problem solving and the interrelations between professional nursing and the women's movement, and a two-term senior integrative seminar for women's studies majors. Although specific applications of the fourth approach may vary in design and scope, they share an emphasis on the direct utilization of findings by students, teachers, and program planners.

Review of the Literature

For the purposes of this planning study, we have chosen to organize the materials in this subsection chronologically. A straight topical or thematic organization would have the advantage of identifying major issues and methodologies and of addressing discrete aspects of the subject of teaching effectiveness. However, the presence of overlapping objectives and methodologies in the literature, compounded by confusions in terminology and varying degrees of methodological specificity and consistency, make thematic or topical organization difficult from a conceptual point of view.

A chronological organization has the advantage of documenting changes in the focus of women's studies evaluative research, particularly when we consider the time frame in which a specific study was undertaken as opposed to the date of publication (or, in the case of doctoral research, the year the degree was awarded).

Women's studies is a field in the process of evolution. This fact is reflected nowhere more clearly than in the literature on its evaluation. Our review of this literature cannot be complete given that many studies that undoubtedly exist have not been reported or listed in the conventional indices. Our review is as complete as it is thanks to the help we received from individual women's studies instructors and programs.

Chronological Review of 1972-75 Studies

By surveying 125 instructors listed in the 1970 Modern Language Association Current Guide to Female Studies, Yamoor et al. (1973) found that, of the 44 percent who responded, only 16 percent reported that their courses were formally evaluated. The team concluded that, "feedback indicates a highly favorable reaction from students, but instructors lacked the

time and research competence to conduct evaluation yielding objective data."

Yamoor et al. assessed the outcomes of "Life Styles of Educated Women" at the University of Minnesota, using a questionnaire they designed (the Minnesota Women's Scale, 1973). The team did not seek to measure changes in attitude, and they had no control group. Therefore, it is difficult to know in what ways the women's studies students in the course differed from other students or, with the exception of the positive course ratings, whether the "favorable" outcomes reported were the result of having taken the course.

Impact of 1973 Wesleyan Conferences. In 1973, two conferences were held at Wesleyan University (Connecticut) to discuss the conceptual and ethical issues involved in evaluative research in women's studies and the relative merits of several evaluation models. As reported by Tobias et al. (1973), among the issues discussed at the first conference was the subject and control of any evaluative research. Also addressed was the appropriate mechanism for measuring the attainment of goals that range from affecting individual values (increasing self-esteem and intellectual competence) and group values (positive attitudes toward women and the ability to work cooperatively), through challenging myths about women and generating new theoretical models within the various disciplines, to instigating institutional and societal change. The conference was important for its attempt to define values and goals for women's studies and women's studies evaluative research on a national level.

As a direct result of the two conferences, several evaluative studies were undertaken using variations on the MAUT model introduced by the late Marcia Guttentag. Although Guttentag's model was in its developmental stage when presented to the Wesleyan conferences, it is important to understand why it

was attractive to the participants, as well as the rationale behind its specific applications.

At the conferences, emphasis was given to the kind of evaluative research that would produce information of direct use to women's studies teachers and students in improving courses and programs. As reported by Tobias et al., MAUT offered the opportunity for course or program participants to develop a range of goals to be achieved, to rank the goals in order of priority and probability of achievement, and to use existing instruments of measuring change or to design new measures to answer questions of specific concern to women's studies teaching and evaluation. In sum, the decision-theoretic model offered the chance of stating both short- and long-range goals and the possibility for women's studies practitioners to control the process and direction of evaluation. (A more detailed discussion of this model is presented in the following subsection.)

One evaluation of a women's studies program that came out of the Wesleyan conferences was conducted by Mangione and Wiersma (1974) at the University of Massachusetts/Boston. Their research was intended both as a descriptive review of the program's educational value and as a planning study for future development. Students, faculty, and staff were involved in setting the objectives to be measured, based on the values discussed at the Wesleyan conferences, supplemented by goals that might be specific to an urban working-student population. The objectives were measured using an interview questionnaire that sought to determine the extent to which the expectations of a variety of students (men, women, and minorities) were being met by the women's studies courses in the program.

Mangione and Wiersma established that, on the whole, student expectations were being met. The researchers noted a synergetic effect from taking more than one course; that is, the more courses taken, the

greater the achievement of goals. Significantly, when asked to rate women's studies courses with respect to their other college courses, 45 percent of the study participants considered them "better" or "much better." Women rated the courses higher than did men, and women's studies "concentrators" rated the courses the highest of all.

Although Mangione and Wiersma made specific recommendations for addressing the needs of men and minorities, for the purposes of this review, the remarks of Ferguson in the introduction of Mangione and Wiersma's report are more relevant. According to Ferguson, "attitude changes" were not assessed because there was no scale sensitive enough to address the issue of consciousness raising in an environment having multiple influences. Moreover, one of the most important findings of the study was "the value students give to cognitive learning in Women's Studies." Being presented with "solid information" and learning techniques for acquiring knowledge and for "assessing their own experiences were more important than attitudinal changes." Presumably, this appreciation for cognitive learning was expressed in the student interviews.

It is noteworthy that, although the acquisition of knowledge and "intellectual competence" were among the values described at the Wesleyan conferences (Tobias et al., 1973), the majority of the experimental studies have sought to measure changes in attitude, using one or more of the instruments developed in the early 1970's to measure the acquisition of new attitudes as the result of the rise of interest in feminism and the women's liberation movement. Such instruments include the Spence and Helmreich Attitude Toward Women Scale (AWS, 1972), the Bem Sex-Role Inventory (1974), and the revised Minnesota Women's Scale (1974). Older instruments, such as the Tennessee Self-Concept Scale (1965) and the Broverman Sex-Role Stereotype Scale (1968), will be noted in place. As Ferguson suggests, some of the questions

that surround when and what to measure in women's studies involve the appropriateness of these scales in an evolving field.

Focus on Sex-Role Attitudes. Ruble et al. (1975) conducted two field studies at the University of California, Los Angeles in which they tested the extent to which three women's studies courses were effective in changing sex-role attitudes of participants in relation to a control group. Five general areas were analyzed: future plans; dislike and distrust of women; agreement with traditional roles of women; nonstereotypic beliefs about sex roles; and perception of sex discrimination. The authors found that women's studies classes significantly increase students' nonstereotypic beliefs and their perception of sex discrimination, while decreasing their agreement with the traditional role of women, even in a one-term course. However, the research found no significant change in career plans or with regard to dislike and distrust of women. The authors also found that women who initially had more traditional attitudes changed the most. Their research could not determine which aspects of the courses accounted for the specific outcomes.

In an unpublished doctoral dissertation, Speizer (Boston University, 1975) reports on her use of the Attitude Toward Women Scale. This 55-item test of profeminist (defined as "liberal") attitudes toward women was used to measure changes in feminist orientation among students who participated in a one-semester course entitled "Perspective on Women in American Society" at Garland Junior College (Massachusetts) in 1974. Speizer found that the scores of the students in that course rose significantly in comparison with the scores of the other women enrolled in the college that semester who were tested. This reflected, Speizer concluded, a heightened awareness of sex-role stereotyping attitudes toward women as a direct result of the course.

Using Rosencrantz's theory of self-concept (1968), Speizer hypothesized that a negative self-concept would change as the result of exploring sex-role stereotyping in a women's studies course. The verbal and written statements of the students, as reflected in class discussions and in journals, were analyzed for units and themes that would express positive or negative feminine self-concept, positive male self-concept, or no self-concept. Speizer's hypothesis could neither be confirmed nor refuted. She discovered that, although the students were willing to talk about women in general, they were reluctant to talk or write about themselves.

Speizer also concluded that different teaching styles did not affect outcomes in the two women's studies courses studied. She comments on the extent to which her findings can be generalized, and, for the purposes of this review, makes a telling point about the AWS:

> Attitudes toward women and men's roles in our society seem to be changing very fast. It is difficult, therefore, to find an instrument which is valid and reliable even within a year of its publication.

Value of Sharing Common Experiences. Borod (Case Western Reserve University, 1975) hypothesized that female participants in her "Psychology of Women and the Acquisition of Sex Differences" course would decrease the discrepancy between their perceptions of typical masculine and feminine characteristics, become less stereotypically feminine in their real and ideal self-perceptions, and become more liberal in attitudes toward women's rights, roles, and privileges. Using the revised Broverman Sex-Role Stereotyping Questionnaire (1972), the AWS, and the Women's Liberation Movement Questionnarie, Borod conducted pre- and post-course testing with followup measurement 3-1/2 months later on both the experimental and control groups. Data analysis did not

support any of her hypotheses. Surprisingly, however, the women who formed the control group became significantly more profeminist in their attitudes toward women and less stereotypic in their perceptions of real self and of the differences between males and females, but only over the period of the study.

Borod, undoubtedly dispirited by her findings, explained them variously. It was possible, she suggests, that the women's studies course might "reinforce, rather than liberalize, sex-role stereotypes," particularly since the course readings and lectures emphasized "researched descriptions of sex differences, and their acquisition." This is an important analysis because it is quite possible that courses in which the materials run counter to the objectives may present special problems in teaching technique (see Hoffman, 1972).

A major implication of her findings, Borod notes, is that, <u>if</u> the goal of women's studies is to change sex roles, then instructional formats that encourage sharing common experiences, as in a consciousness-raising group, might be more effective. Borod also suggests that exposure to the milieu created by the women's movement and participation in the study itself may have influenced the scores of the control group.

<u>Further Study of Attitude Changes</u>. Shueman and Sedlacek (1976) in 1975 conducted a study that combined the legitimizing aspects of the descriptive review with experimental research on attitude changes. The authors compared students in nine women's studies courses with students in an undergraduate psychology course at the University of Maryland. In addition to gathering demographic information, data on attitudes relating to career goals and women's issues, and course evaluations of teaching effectiveness, they used the Bem Sex-Role Inventory (1974) as a measure of psychological androgyny and the Tennessee Self-Concept Scale (1965) to assess changes in feminine stereotyping and self-concept as the result of the courses.

The researchers made an effort to align the concepts they wished to measure with information on the objectives of the women's studies program outlined in program brochures.

The authors reported that, in 1975, women's studies courses were largely of interest to women, were taught mainly by women, were offered in the humanities and social sciences, were perceived by the students as academically rigorous, and did increase student awareness of sex roles and the "situation" of women in society. They found, however, no significant change in self-concept or stereotyping that could be attributed to the courses, and speculated that one semester may be "too short a time for any significant change to occur and be reflected in a standardized measure."

Scott et al. (1977), reporting on research they conducted in 1974, found that some student attitudes can be altered by a one-semester or one-quarter course in women's studies as measured on the AWS. Their subjects were students in two midwestern colleges. The authors suggest that the influential factors in liberalizing scores were the relevance of the materials in the course to a predominantly female enrollment, emphasis on student involvement in discussions, and the use of strong historical women as role models, as well as the contemporary role models presented by highly competent guest speakers.

In an evaluation of a course entitled "Achievement Motivation in Women: Psychological and Sociological Perspectives" (1975) at the University of Minnesota, pre- and post-course assessment on four scales (Self-Assessment Scales, Attitudes Toward Women as Managers, AWS, and Achievement Motivation Inventory) did not detect any changes that could be attributed to participation in the course. In student self-evaluations, however, many reported that they had not only learned new facts, theories, and approaches, but had also gained new perspectives on themselves as

women and as scholars and were much more ready to assume responsibility for their educations.

Thames (University of California, Los Angeles, 1975) reported in an unpublished doctoral dissertation that the students who scored the highest in feminism on the AWS were those enrolled in women's studies courses in which the instructor announced either verbally or through course objectives that she was a feminist. In courses in which the instructor was explicitly or implicitly "critical" of women, the scores on the scale were significantly lower. Thames called for the development of a questionnaire that would tap the most feminist and radical interests and attitudes of students in order to measure the full spectrum of responses to women's studies.

The MAUT Model

Although Guttentag's Multi-Attribute Utility Decision-Theoretic model did not reach a national women's studies readership until 1978, versions of MAUT were published as early as 1973. And Guttentag's presentation at the 1973 Wesleyan conferences inspired several studies based on her methodology, including that conducted by Brush et al. during 1974-75 at Wesleyan University. The results of part of that study were circulated as early as 1974, and a final version, "The Paradox of Intention and Effect: A Women's Studies Course," received wider distribution in Signs (summer 1978).

Report of Guttentag et al. In "Evaluating Women's Studies: A Decision-Theoretic Approach" (Guttentag et al., 1978), the authors characterize some of the problems that traditional research encounters when addressing itself to women's studies. First, formal evaluation, usually conducted by an outside evaluator, may not ask the kinds of questions that most women's studies practitioners want answered. These questions deal less with success and more with ways in which courses and programs can be improved.

Second, the traditional models of evaluation often measure goals that are not shared by the various groups that have an interest in the project. Third, goals and expectations may change during the period of evaluation. The traditional models, which are oriented toward summative results, may not be able to accommodate the revision.

MAUT involves the participants of the evaluation directly in setting goals and in establishing priorities that are specific to a particular course (or to a number of courses in a women's studies program), as well as in assessing the probability of achievement. The evaluators are responsible for finding or developing techniques to measure the degree to which any given goal is met, and for collecting and analyzing the data using Bayesian statistics. The opportunity for different groups to establish different goals for the evaluation and to assess the probability of achievement seems particularly important in a field in which the goals range from increased intellectual self-confidence to changing society.

An important part of the process involves identifying the areas to be evaluated and making the definitions operational; that is, breaking the expected outcomes down to specific attitudes or actions and identifying the alternative instructional formats for implementing the goals.

The MAUT model, the authors argue, permits the kind of "iteration" and flexibility needed in evaluating the complex expectations of women's studies courses and programs. Their presentation does not specify instruments that are compatible with MAUT; they do say that priority should be given to developing new and appropriate techniques for measurement. Analyzing the data collected at any point in the process, assigning "importance weights," and converting these weights mathematically by Bayesian statistical analysis permits a series of interactions between probability and testing results so that a

judgment may be made about which sort of instructional mode or course best achieves the desired goal.

Report of Brush et al. Brush et al. begin their report of a specific application of MAUT with a statement that needs to be viewed in the context of the evolution of women's studies. Of the two sets of goals that are emphasized by practitioners -- "the traditional academic goals of intellectual mastery of subject matter and the imparting of a substantial amount of information" and the "less traditional goals of personal change" -- the researchers considered it more worthwhile to test whether the second set of goals, "resocialization," might be achieved in a classroom setting during a one-semester interdisciplinary course. The goals of resocialization, as specified by the participants, were categorized by the researchers as improved self-esteem, improved self-concept, and commitment to a "feminist ideology," defined as changed attitudes toward the social role of women, raised career aspirations, and consciousness of sex bias.

The instruments used included the Bem Sex-Role Inventory, the Broverman Sex-Role Stereotype Scale, the I Am test (Kuhn and McPartland, 1954), and a revised version of the Minnesota Women's Scale, supplemented by questions to elicit information on career plans and reactions to sex bias that was used to "graph ideology." When the results of the Minnesota questionnaire were discussed a year later with the students, "spontaneously mentioned changes" attributed to the women's studies course were recorded and analyzed.

As suggested by the title of the Brush et al. report, the authors found a paradox of intention and effect on several levels. The first paradox involved the students and the course goals. Consistent with the findings of Speizer (1975) and Ruble et al. (1975), entry level scores of the women's studies enrollees initially were high. In the Brush et al.

study, the scores of some of their control groups were also high. (The number of men in the study was small, and their scores were not tabulated.)

Brush et al. found that women who came into the course with more traditional attitudes changed the most. For the rest, the course tended to validate or clarify prior beliefs and to reinforce their retention by providing information, forms of argumentation, and the opportunity to articulate belief, information, and critical analysis. In at least one of the authors' control groups, initially strong profeminist beliefs weakened over time, presumably because they had not received the kind of reinforcement and strengthening that a women's studies course can provide.

The second paradox involved the measurement instruments. For all the authors' effort to choose sophisticated instruments, the statistical measures failed to detect the kinds of changes that the students reported in their interviews a year or more after the completion of the course. Content analysis of interview data suggested some changes in self-concept, and particularly in the areas of heightened self-confidence, in the value placed on independence, in the acceptance of qualities of masculinity and feminity about which some of the women's studies students initially had been ambivalent, and, on the part of 6 of the 16 women interviewed from the 1974 classes, in an increased respect for women.

Of the various conclusions the authors drew, perhaps the most relevant for this part of the literature review is their assessment of the changing goals of this particular women's studies course since its inception in 1970. Initially, changes in behavior and attitude -- the second set of goals -- were of primary concern. By 1978, however, when the study was published, both the instructor and the students stressed "intellectual mastery" as their primary objective in giving or taking the course. The instructional format had also undergone unspecified changes.

Chronological Review of 1975-78 Studies

Interviews conducted in 1975 by Bose et al. (1977) of women's studies majors and nonmajors at the University of Washington support the findings of Brush et al. -- i.e., that students report positive changes in self-image as a result of their women's studies experiences. The students in the Bose et al. sample also reported "an increased awareness of their own needs, and more faith in their ability to fulfill these needs," thus confirming the value (or perhaps the necessity) of independence.

Effects of Introductory Sequence. A study of the effects of the introductory sequence at the University of Minnesota (1976) used the Washington University Sentence Completion Test of Ego Development (1970) and the AWS. No pre- to post-course changes in ego levels or attitudes were found either in the control groups or among the students who had completed the women's studies introductory sequence. The research established, however, a correlation between the women in both groups who had more liberal attitudes toward women and a higher than average ego level. The report suggests that this correlation may stem from the ideals of self-realization and identity supported by the women's movement.

Coffman (1978) reported on her 1976 investigation of the effect of introductory women's studies classes on sex-role stereotyping. Coffman hypothesized that these classes would reduce sex-role stereotyping by increasing androgyny and masculinity scores on the Bem Sex-Role Inventory and that the control group of introductory sociology students who expressed an interest in taking women's studies courses would not. Coffman also sought construct validity for psychological androgyny, which was defined operationally as the ability to respond flexibly to situations, as opposed to reacting according to sex-role prescription. She anticipated increased masculinity scores as

a preliminary step to androgyny based on the findings of Steiger (1977) and others.

Coffman's data analysis did not support her hypotheses with respect to the women's studies introductory classes: neither these nor the introductory sociology courses significantly reduced sex-role stereotyping, which was a goal of the women studies course at the University of Washington.

Two other studies conducted during this period were exploratory in design and relied heavily to exclusively on student self-report. These studies addressed, respectively, program effectiveness (Talburtt et al., 1977) and impact of women's studies courses on students' lives (Elovson and Cockcroft, 1977). Both studies illuminate some of the cognitive and pedagogical issues that are revealed in the three case studies discussed later in this section.

Study of Program Effectiveness. The goal of Project WELD (Women's Education: Learning and Doing) was to evaluate the impact, outcomes, and implications of three women's educational programs: internships, defined as career exposure in out-of-classroom settings; women's studies academic courses (not necessarily a part of a formal women's studies program); and skills development classes or workshops, including assertiveness training, career planning, and personal growth experiences. The programs were housed in eight institutions, six of which were primarily women's colleges and two of which were coeducational universities with women's centers. It should be noted that although the programs were viewed separately, any combination of the three may be included in a women's studies program.

Using a pre-tested form, Talburtt et al. interviewed students (most of whom had completed the programs a year prior to the interview), teachers, administrators, and supervisors, and they consulted

with a national advisory board. The interviews provided both quantitative and qualitative information, and primacy was given to returning the results to the institutions to document and/or to improve their programs' effects.

Talburtt et al. found that, for the most part, the students chose the various programs to gain the general outcomes promised, that the outcomes were achieved, and that unexpected results were achieved that were highly valued by the students. Internships promoted professional skills and career exposures, but they also increased self-confidence and sense of personal potential. Skills classes were reported as developing interpersonal skills. They also developed self-confidence, independence, feminist perspectives, and self-understanding. In other words, each type of program accomplished its objectives and fostered outcomes that were unanticipated. When the students compared their programs to their other classes, in no case did they report that the "traditional curriculum" developed qualities that they considered important in a significantly better way.

As the authors note, program structure has impact on educational outcomes, for the outcomes were indeed different, even with some overlappings.

> Women's studies would appear to be the most direct route to fostering feminist consciousness about women's past and present. An internship encourages the development of professional potential and antidiscrimination skills better than either [of] the other two program types or the traditional curriculum. Yet with the exception of a sense of women's historical past and professional qualities, the skills experience seems to promote every other attribute better than the traditional college experience.

Internships and skills classes may already be a part of a women's studies program (as they are at Portland State University, for instance) or may be a part of a women's studies course that differs from the ones in Talburtt's study. Nevertheless, the findings of this research have implications for building curriculum in and out of women's studies, particularly in an era in which students are acutely sensitive to the relation between their education and future employment (Bose et al., 1977).

While a majority of the women's studies students found the course impact to be greater than other educational experiences, 23 percent of this group rated their classes as having the "same" or "less" impact, a higher percentage than in either the skills or the internship groups. Satisfaction with the "academic content" was slightly higher in women's studies than in the other two programs. The authors propose that, in this study, the comparability between women's studies and traditional courses made it easier for the students to know what to expect by way of cognitive development and hence were better able to measure the cognitive development gained. This study also found that teachers slightly overestimated the possible outcomes of the women's studies courses, particularly in the areas of creativity and leadership.

Impact on Student's Lives. The exploratory research of Elovson and Cockcroft (1977) was designed to generate a broad base of data to illuminate curricular and pedagogical issues and to develop a "reliable and sensitive" instrument of assessment. In examining the impact of participation in women's studies classes on students' lives over time, the authors proposed an alternative to the "defined-objective" approach to evaluation. To avoid the methodological difficulties, the "paradox of intention and effect," and the unanticipated outcomes that so often crop up in the literature, the authors

recommended the "empirical-inductive" approach and the expanded use of student self-reporting.

Using a combination of open-ended and directed-attention questions, Elovson and Cockcroft gathered "impact" data from some 200 women's studies students. Ninety-three percent of the students reported impact. When individual impact scores were correlated with the number of women's studies courses taken, it was found that the more courses taken, the greater the number of scores of support and reinforcement for previously held ideas, feelings, and behaviors, thus arguing for the internalization of impact in the students' lives.

Ninety-five percent of the students responded that they had encountered new material in their women's studies courses. Moreover, 19 percent reported that the new information affected how they viewed issues outside women's studies. The students also reported that the female teachers had served as role models in competence, energy, intelligence, achievement, and commitment to scholarship, as well as in personal warmth and expression.

Three Case Studies

The case study approach to evaluative research in women's studies is perhaps the most widely represented in the informal literature. The examples presented here combine traditional course ratings with an emphasis on instructor monitoring and analysis of course progress, leading to changes in pedagogical approach, content, and class structure. These studies differ in the degree to which they also contain aspects of experimental design and descriptive review. They are joined, however, by the degree of attention given to the content, structure, and teaching strategies that make operational course goals.

Research of Davis et al. The research of Davis et al. (1975-76) encompassed evaluating the success of the "Images of Women in Literature" course in

achieving instructional goals, measuring changes on the AWS, comparing demographic characteristics with a control group, gathering feedback on instructional effort during the term, and documenting the suitability of the course as a permanent offering.

The authors found that the objectives of the course were met. The students gained a greater knowledge and appreciation of women. Post-course scores on the AWS showed significant improvement as the result of instructional effort, while the scores in the control group did not improve.

Research of Kritek and Glass. Kritek and Glass (1978) evaluated a course offered in a professional nursing program and cross-listed with the women's studies program at the University of Wisconsin/Milwaukee, thus drawing students from each. The course was oriented toward the interrelations between nursing as a profession and the women's movement, and it emphasized creative problem-solving.

The authors found that the examinations and projects assigned encouraged originality, initiative, and self-direction, and fostered a feminist perspective in nursing. Students reported that class discussions had led to on-the-job assertiveness with respect to the sexist discrimination they perceived.

Kritek and Glass used the Allport Study of Values, an evaluation measure of their own design, and student ratings of materials and instructors. All measures showed significant progress toward obtaining course objectives.

Research of Register et al. Register et al. (1978) describe at length the design, progress, and evaluation of a two-term senior integrative seminar at the University of Minnesota. The purpose of the research was to determine how best to meet a women's studies program requirement that majors have the opportunity "to direct knowledge and methodologies

gained from other courses and experiences toward topics of an interdisciplinary nature with a focus on women."

During the first term of the course, the four instructors worked as a cooperative team, alternating in presenting lectures on the following topics: assumptions that characterize scientific research and analysis of the process by which theories are constructed; anthropological approaches, including cross-cultural comparisons; literary criticism; Jungian psychology; Marxist and Marxist-feminist methodology; and biases and problems in the behavioral and social sciences. The lectures and subsequent discussions were intended as overviews and preparation for the second term's requirement that students develop their own research topics on the theme of women, culture, and power.

The second term was structured to give students the opportunity to work on their research and to meet once a week for student-led discussions of a work of feminist theory. The final 2 weeks were devoted to the oral presentations of student projects before these were submitted in written form.

A summary of the evaluation cannot do justice to the complexity of the task or the results. At the end of both terms, students were asked to evaluate such matters as: the difficulty of the reading and its appropriateness both to an interdisciplinary approach and to their own feminist concerns; the success of the teachers in presenting subject matter and as human beings in the classroom; and their own and other's participation in the seminar.

Register et al. began their discussion of the results by saying that the "process of self and mutual evaluation has not always been pleasant, but we feel that it is a vital part of the growth of women's studies." Students and teachers had entered the course with a high level of expectation of themselves

and one another. The conflict between the academic emphasis of the teachers and the action-orientation of some of the students, as well as the teachers' own struggles to integrate academic values and feminist action, permeated the course. Students and teachers had tried to clarify their own understandings of such terms as "feminist" and "interdisciplinary," but the diversity of the perspectives resulted in shifts of focus: the center did not always hold.

The instructors were surprised at the levels of difficulty the students reported in the readings. They were equally surprised that what they had perceived as an experiment in cooperative feminist teaching was viewed by a number of the students as a class top-heavy with instructors who formed a silencing block of authority. The instructors' assumption that mutual trust and respect had been developed before the beginning of the course proved to be unfounded.

The instructors also discovered the "real risks involved in attempts to integrate the academic, the political, and the personal and to be teachers, allies, and friends." Some of the criticisms the students had of the course -- the abstractness of the overviews, the relative lack of student input into the design and progression of the course -- became personally directed toward some of the instructors. Ironically, the encouragement to make a radical critique of education often finds the encouragers on the receiving end of the stick. Yet for all the silences, disappointments, and crosscurrents of feeling, the students on the whole rated the course, the teachers, and their own learning positively.

The findings of this evaluation led to modifications in the course requirement, particularly in the direction of a realistic reassessment of what an integrative course promises and what it can actually deliver, and in the development of options for meeting the requirement.

Summary of Findings

This review of the literature indicates that the standardized measures used to gauge the effectiveness of women's studies teaching have not proved to be sensitive to the changes in perception and attitudes that are reported in student interviews and in the exploratory and case study research. The exception to this is the Attitude Toward Women Scale, the workhorse of experimental design research.

In the studies that reported using the AWS, four found pre- to post-course "liberalization" of scores among women's studies students and three did not. The questionnaire used by Ruble et al. (1975) detected statistically significant changes in attitude toward the traditional role of women, a reduction in stereotypic beliefs, and an increased perception of sex discrimination on the part of students as a result of taking a women's studies course. It uncovered no changes in dislike or distrust of women or in career plans stemming from the course. Shueman and Sedlacek (1976) found increased "awareness" of sex roles and the "situation" of women as the result of taking women's studies courses.

Several of the researchers have reported that the greatest changes occur among women students who have the most traditional attitudes. Brush et al. (1978) found that, for the rest in their sample, the women's studies course served to clarify and strengthen prior beliefs and to equip them with the information and intellectual skills to use the course effectively.

On the whole, the research that utilized student self-reporting and student ratings found that women's studies courses achieve their goals and that students judge these courses as having a powerful impact on their lives and educations.

It is a major question to what extent the bulk of the research reported to date -- experimental design

studies -- bears on the issues of cognitive learning, and hence on how comprehensively the issues of teaching effectiveness in women's studies have been addressed. With the exceptions of the attention paid to the development of cognitive skills and the acquisition of cognitive content in the exploratory and case study approaches, and of the report of student appreciation of cognitive learning in Mangione and Wiersma (1974), the literature of evaluative research in women's studies tends to be silent on the traditional academic goal of "intellectual mastery."

Finally, chronological review of the literature indicates that an early trend toward assessing effectiveness almost exclusively in terms of attitude change has been superseded by a more multivariate approach. In this respect, the work of Talburtt et al. (1977), Elovson and Cockcroft (1977), Davis et al. (1975-76), Kritek and Glass (1978), and Register et al. (1978) is exemplary. The Talburtt and Register studies are also noteworthy for their inclusion of teacher self-evaluation.

Women's Studies Values and Teaching: Theory and Practice in the Classroom

In section 2, we discussed the process-product paradigm (Doyle, 1975), which favors a bidirectional and multivariate model for assessing teaching effectiveness. The accounts of women's studies classrooms in this subsection are presented for the purpose of conceptualizing some of the variables and interactive modes that obtain in women's studies teaching and learning as described by women's studies teachers. Issues of importance to future development, and hence to the future evaluation of women's studies, are also raised.

The Participant-Centered Classroom

The theory behind the kind of women's studies teaching that places the participants at the center

of the course argues that the most effective learning is active engagement with the materials and other students in the course. It argues also that old, passive socialization and thought patterns can be challenged and new patterns established through the direct participation of the student in the learning process.

Howe (1970) describes a writing course for women in which she attempted to break the patterns of assumed inferiority and dependency on the teacher as a figure of authority; that is, to overcome the barriers of intellectual domination and trivialization women students often have encountered in their past education. Howe faced squarely the issues of intrapsychic and interpersonal conflict engendered by the process. It is, she argued, a long step toward autonomy and power for the student and the teacher both to risk the anger and bewilderment of disruption, for what is being reconstructed is not just students and their writing but the classroom itself.

Hoffman (1972), in describing a women's poetry course, argued for locating the process of explicating texts in the students themselves, as a means of counteracting the negative content often found in the literature. Hoffman viewed students and teachers reading together as a social act that in itself projects a community different from the patterns of isolation and despair found in some of the women writers. Martyna (1976) suggests that, in a course that explores the ways in which men and women use language, the interactive classroom is an ideal arena for students to observe and begin to change their own thought and speaking patterns.

The "talking classroom," however, is not without its attendant discomforts. Not only are discussions heated and personal (Talburtt et al., 1977), they often do not stick to the subject, a phenomenon of particular, although by no means exclusive, discomfort to the teacher. Stizel (1977) reminds us that the classroom is only <u>one</u> component of a course. Dis-

cussions, even when apparently irrelevant, can "revitalize memories, alert readers to mindsets that structure responses, and raise and focus issues that become contexts" for the reading, writing, and discussion of books that go on outside the class.

Proponents of the participant-centered classroom see it as the structured space in which teachers and students learn together to connect personal experience to the substance of books and to the larger issues and analyses that lie outside the classroom. This connection, particularly in the context of developing mutual trust and respect for women, is seen as a value in itself and as a means of ensuring learning that has a context (Howe, 1977).

Not all women's studies teachers and students, however, rapturously embrace the personal and open-ended classroom discussion. Fireman (1973) noted that the "emotional element" that she found running high in the history of women course she taught made concentrating on the subject matter problematic. Kritek and Glass (1978) reported that some of their students were disconcerted by what they perceived as a lack of structure (authority) in that part of the course in which they were asked to take greater initiative for their own learning. Davis (1978) described a protracted struggle in a research seminar between the students' request to be "mothered" during the process and her own conflicting impulses to mother and to demand student independence. Such accounts raise two issues that have not been systematically investigated: the prior training and background of students and teachers in women's studies, and the contexts from which they work.

Among those women's studies teachers whose backgrounds, pedagogical preferences, and course objectives favor the participant-centered classroom, there is growing awareness that certain goals do not just happen spontaneously. These include: mutual respect, trust, and community; a shared leadership

between students and teachers; cooperative projects; integration of affective and cognitive learning; and the integration of theory and action. Translating goals from rhetoric to teaching and learning involves building skills and dedication in the classroom and a continued understanding of why this is important (Schneiderwind, 1978-79).

The Content-Centered Classroom

The content-centered classroom is not necessarily synonymous with the lecture-centered classroom, nor does it necessarily imply that the teacher is the center of the course. Values and objectives may overlap with those of the participant-centered classroom, particularly the connections made between personal experience and subject matter. As Stoper observes, the emotional engagement of teachers and students in the content and processes of any course in women's studies must be combined with "sensitivity" in structuring classes to make the emotional engagement a "powerful intellectual asset."

Lectures, however, are a necessity in many large courses, and they are often viewed as the most effective means for concisely conveying large and complex bodies of knowledge, as well as for demonstrating paradigm-challenging and feminist forms of argumentation and approaches to content (Wolf, 1978). The lecture format may also be seen as a vehicle for presenting strong role models for women students, particularly with respect to scholarship and the professions. Finally, from the perspective of the responsibilities that women's studies teachers feel for both subject matter and students, the lecture may be seen as a format that clearly delineates the function of the teacher as a person who cannot solve specific problems of decisionmaking in students' lives, but who can offer the broad context and analysis of where inequalities and oppression come from in institutions, culture, and society (Eileenchild, 1979).

Because the lecture-discussion format is traditional in higher education, it has received less attention in the published and informal literature on women's studies teaching. Essays in the Sociologists for Women in Society Newsletter (July 1978), however, address feminist, interdisciplinary, and pedagogical perspectives relevant to the lecture-discussion format. Of particular note are the essays that describe the objectives of developing the "conceptual tools necessary to articulate [personal] understanding in structural terms" of discussion grounded in course lectures and readings (Gould, p. 4); of helping students to identify the ways in which they contribute to the maintenance of sexual inequality; and of encouraging equal participation even in large lecture classes (Thorne and Parrington, pp. 10-11).

Transitions in Women's Studies Classrooms

Stoper speaks to several issues that are important variables in understanding the changing objectives that might occur in women's studies classrooms. She describes, for instance, the rationale behind abandoning small group discussions that lacked the intimate, voluntary setting essential for their success. Stoper also reports that the growing body of findings on the status of women in her field of political science has led her to adopt a less interdisciplinary approach in her political science women's studies course. On the other hand, the growing mass of material in other areas as well has led her to participate in a team-taught, interdisciplinary course in non-American history, literature, and philosophy that was enthusiastically received.

Course objectives may change for other reasons. As the field develops and programs have the resources to expand and differentiate offerings, courses and their objectives become distinguished by kind or level of specialization. And as some of the literature reviewed in the previous subsection suggests, consciousness has already been raised and sex-role

attitudes and stereotyping changed, at least among some women's studies students and particularly among students who become majors or concentrators.

Finally, some of the published and informal literature on women's studies teaching raises issues that bear upon future evaluation of teaching effectiveness in this evolving field. Whether changed consciousness occurs before a student takes a women's studies course, during that course, or as a delayed reaction to the course, teachers face helping students with "transformed values" to live and work productively in a society whose values have not undergone radical transformation.

To paraphrase Morgan (1978), teachers cannot alienate students from their culture and leave them there, but must use the accumulated knowledge of interdisciplinary studies to help students in gaining a better understanding of the process of alienation and of how to deal with it. Hersh (1979) advocates teaching the sociology and psychology of role change as part of the course or curriculum, and Morgan suggests examining the historical behavior of oppressed groups and the dynamics of individuals and society in transition. Arpad and Arpad (1979) propose building into a women's studies course an examination of consciousness itself as a way of helping students mediate between their own past and present identities. They suggest that this be done by reviving and revitalizing the old pedagogical strategy of philosophic inquiry into the nature of things.

Summary

Our overall impression from the published and informal literature is that women's studies is feminist, at least in this sense: women's studies students and teachers strive in the ideal to make their classes a mediation between what has been and is and what is and what might be. This process of simultaneously studying the past and present and

making a future characterizes the state of women's studies knowledge itself. In this evolutionary and hypothetical stage, it is clear that any evaluation of teaching effectiveness in women's studies must take into account the multivariate, changing, and bidirectional nature of the field. Students affect teachers, teachers affect students, and both affect the state of knowledge itself.

5. CONCLUSIONS AND RECOMMENDATIONS

In our conclusions and recommendations, we propose both long- and short-term goals for research on teaching in women's studies and for the evaluation and improvement of teaching effectiveness.

Research on Teaching

Women's studies is a major innovation and reform in education, and women's studies teaching deserves both basic and applied research. Research that is directed toward understanding the relationships between variables that obtain in teaching and learning would provide a knowledge base from which research that is applied, problem-solving, and evaluative could optimally develop.

As stated in section 1, women's studies courses and programs issue from several different educational traditions have differing or multiple objectives. They are further complicated, from the perspective of research, by the overlay of feminism and the interdisciplinary nature of some of the courses.

Given these considerations, we recommend that further research on teaching take into account the complex variables and paradigms outlined in section 2. To our knowledge, research based on the multivariate and bidirectional models has not addressed itself specifically to women's studies courses or programs. Many crucial questions about the relationship between effective teaching in women's studies and effective teaching in both traditional and innovative fields remain to be investigated.

In the short run, the inductive, observational, case study approach, combined with experimental research, seems the most suitable methodology for determining the complex variables in women's studies teaching and learning. Inductive or case study

research has the advantage, as Doyle (1975) points out, of generating hypotheses that can be further refined and tested by experimental research. It is our assessment of the experimental research in women's studies to date that it has not so much failed to ask the right questions, but that the line of questioning has not been comprehensive enough and has been overly focused on attitude change. In addition, measuring devices have not been particularly sensitive.

On the other hand, we found the study of Talburtt et al. very useful in terms of understanding the impact of different kinds of women's educational programs. The impact studies, however, do not necessarily explore the specific dynamics of teaching and learning in any given course or program. As women's studies teachers, we found the case study work of Davis et al., Kritek and Glass, and Register et al. most useful. Because each approach, by itself, has limitations, we favor a combination of inductive/exploratory, case study, and experimental methods in future research on teaching in women's studies.

Furthermore, the experimental researchers whose work did not detect statistically significant mean score changes in attitude after one term of a women's studies course call for longitudinal studies. The research that does not find changes in group scores does find changes in individual scores attributable to having taken a women's studies course. The exploratory research of Talburtt and Elovson determined that women's studies courses have impact over time both on individuals and on groups. We recommend, therefore, that research undertaken to determine cognitive and affective changes in students as the result of taking women's studies courses should assess both individual and group scores and should be longitudinal.

To begin developing a comprehensive and systematic knowledge and data base, we suggest that four

kinds of women's studies courses receive intensive attention in future research, with particular emphasis on cognitive learning:

1. A lower-division introductory course, such as "Introduction to Women's Studies" or "Women in Contemporary Society."

2. An upper-division discipline-centered course aimed primarily at mastery of the subject matter and development of skills of scholarship and critical thinking and analysis; for example, "Psychology of Women," "Women and the Law," or "Earlier American Women Writers."

3. An upper-division interdisciplinary or integrative course, such as "Feminism: Theory and Practice" or "Women in a Cross-Cultural Perspective."

4. A skills acquisition course, such as "Women in Management" or "Assertiveness Training."

In keeping with Biddle's emphasis on the importance of the institutional or community context and Talburtt's emphasis on the different institutional realities of women's studies, we recommend that these four types of courses be systematically examined in five different institutional settings: a large, state-supported university; a middle-sized urban institution (either public or private); a small liberal arts college; and a community college.

In addition to the variables of teaching and learning considered in any research on teaching effectiveness, several issues specific to women's studies need to be considered. These are discussed below.

Feminism and Teaching Effectiveness

A high proportion of the teachers who responded to our survey (see appendix) agree that feminist or feminist-humanist approaches to teaching exist in terms of subject matter, relations in the classroom, and general mindset. In addition, some experimental research has concluded that the presence of a feminist in the classroom does make a difference, particularly for women (Blumenhagen, 1974; Thames, 1975; Coffman, 1978; Farley, 1978). Research should be undertaken to determine the variables of teaching and learning that are influenced by feminist or feminist-humanist approaches in teaching.

Interdisciplinary Approaches And Teaching Effectiveness

We have been unable to locate experimental research that assesses the relationship, if any, between an interdisciplinary approach to subject matter and teaching effectiveness. Baxter (1974) hypothesizes that the interdisciplinary women's studies approach uncovers a conflict-ridden dialectic between women and culture. The structure of the subject (Bruner, 1963) and the organization and design of the subject matter (Davis, 1977) influence learning outcomes. Implications of the interdisciplinary approach for teaching effectiveness in women's studies should be investigated.

Formative Experiences of Women's Studies Teachers and Teaching Effectiveness

Research should be initiated on the formative educational and life experiences of women's studies teachers. We are concerned with the relative impact of the formal academic background of women's studies teachers in an interdisciplinary as opposed to a discipline-specific course, as well as with the relation of these variables to the question of effective teaching. Information of this nature bears

importantly on the development and revision of the women's studies curricula at both the undergraduate and the graduate levels, particularly as curricula are related to the preparation of future teachers.

Sex of Instructor and Student And Teaching Effectiveness

Further research is needed to test the conclusions of McKeachie et al. (1971), Ferber and Huber (1976), Harris (1976), and Kashak (1978) regarding the sex of the instructor, the sex of the student, and sexism within the context of the women's studies classroom. We are also concerned about the extent to which the learning and cognitive styles of females should be considered in the design of research conducted to assess learning outcomes and teaching effectiveness.

In the long run, research must be undertaken to develop appropriate and sensitive measures of learning that can assess multiple course objectives, including changes in attitude, resocialization of relationships in the classroom, mastery of subject matter, critical thinking, acquisition of skills, and the integration of these objectives into effective teaching and learning. As Scott (1975) notes, one of the major problems in research on teaching has been the lack of adequate measures of student learning.

Two additional concerns need to be addressed. Who might do the suggested research? To what uses might results be put?

It is our opinion that research on teaching in women's studies would best be conducted by teams of individuals who come from women's studies and from the various disciplines associated with educational research. We think that initiating and developing working relationships between educational researchers and women's studies practitioners would be of mutual benefit to both; it would enhance the probability

that the important questions are asked and that the results are interpreted and disseminated in an appropriate manner. Women's studies students, teachers, and administrators have much to learn from educational researchers, just as educational researchers have much to learn from women's studies students, teachers, and administrators.

If the object of research is to provide information and understanding about the nature of a particular phenomenon, research findings should not be used for evaluative purposes, other than in the formative mode. Results of research should not be used for advocacy or nonadvocacy in any particular women's studies program. Participants must be guaranteed the protection given other subjects of research.

The research suggested above is of critical importance not only for understanding teaching effectiveness in women's studies, but for understanding more about the education of women in general. The classroom is the heart of educational change, and understanding more about women as teachers and learners will be of use to those concerned with providing equal educational opportunity for women both inside and outside of women's studies.

Evaluation and Improvement of Teaching

As stated in section 3 of this monograph, we agree with the AAUP position paper that multiple measures must be used to evaluate teaching. Ideally, before proceeding with recommendations for the evaluation of teaching, one should wait until sensitive and appropriate measures of student learning are developed and tested and until a knowledge and data base has provided criteria of effective teaching in various kinds of women's studies courses. However, women's studies practitioners are faced with immediate needs for information on which to make personnel, curricular, and program decisions.

Despite the problems with the validity and usefulness of students' ratings described in section 3, 94 percent of our questionnaire respondents indicated that their courses are being regularly evaluated using student rating forms of college or university design, women's studies design, or both. Thirty-nine percent indicated that they are either neutral or dissatisfied with their current evaluation instrument. In this sense, women's studies teachers are in the same predicament as most other teachers in higher education (Cross, 1977).

As a first step, we recommend that funds be made available to those programs that wish to analyze evaluative data collected previously. Women's studies programs are underfunded, and preliminary analysis will not be possible for most programs without outside assistance. Thus, technical assistance should be made available, as necessary, to programs involved in data analysis.

In addition to data analysis, efforts should be made to factor-analyze and refine evaluation instruments. After preliminary work is completed by programs, the National Women's Studies Association should systematically collect and analyze student rating forms from all women's studies programs where such forms are in use, and begin the process of determining the criteria for evaluating effective teaching in various types of women's studies courses. Sample forms should be made available without cost to programs and individuals.

As a second step, we recommend that in-service workshops on the evaluation of teaching be conducted for women's studies students, teachers, and administrators. Workshops should include theoretical issues in the evaluation and improvement of teaching, instrument selection, methods for systematically collecting data (including both student and instructor self-report data), data analysis, and the interpretation and dissemination of results.

Third, we recommend that the participants and leaders of the evaluation workshops develop a resource/guidebook for those individuals who are unable to participate in the workshops. Again, theoretical and methodological issues should be covered, as well as sample evaluation instruments and resources for technical assistance.

Finally, we are in agreement with Davis (1977) that the design and organization of a course is an important variable in effective teaching. In-service course and curriculum design workshops should be made available on a regional basis to women's studies students, teachers, and administrators. The focus should include both discipline-centered and interdisciplinary courses and curricula, with attention given to teaching methodologies and course design.

To this end, we recommend that the National Women's Studies Association develop and maintain a clearinghouse for course outlines and bibliographies. As Gerda Lerner has noted (1979), the willingness of women's studies teachers to share course outlines and bibliographies with each other has advanced the field more rapidly than otherwise would have been possible. We urge that, whenever possible, course outlines and bibliographies be submitted to the clearinghouse with both instructor and student evaluations. Such information would greatly assist individuals in refining and improving course design.

In sum, the relationship between educational research and educational practice must be bidirectional. The progress of research on teaching suggested in the previous subsection depends on both the information and insights of women's studies practitioners. If evaluative information is collected and analyzed in a systematic manner, and if opportunities are available for improving course design and teaching methods, the probability of maintaining and improving effective teaching will be enhanced in any particular program. Further, the information provided

by practitioners will ensure that researchers direct their lines of questioning to those issues and concerns that are of importance to practitioners.

BIBLIOGRAPHY

American Association of University Professors Committee on Teaching, Research and Publication. "Statement on Teaching Evaluation." AAUP Bulletin, 1974, 60 (2), 166-170.

Arpad, S., and Arpad, J. "Consciousness Changing in the Women's Studies Classroom." Proceedings of the Conference in Communication, Language, and Sex. Massachusetts: Newbury House, 1979.

Baxter, A. "Women's Studies and American Studies: The Uses of the Interdisciplinary." American Quarterly, 1974, 26, 433-439.

Bergquist, W.J., and Phillips, S.R. A Handbook for Faculty Development. Vol. I. Washington, D.C.: Council for the Advancement of Small Colleges, 1975.

Biddle, B.J. "The Integration of Teacher Effectiveness Research." In B.J. Biddle and W.J. Ellena (Eds.), Contemporary Research on Teacher Effectiveness. New York: Holt, Rinehart, and Winston, 1964.

Blumenhagen, K. "The Relationship Between Female Identity and Feminism." Doctoral dissertation, Washington University, 1974.

Bonneparth, E. "Evaluating Women's Studies: Academic Theory and Practice." In K. Blumenhagen and W. Johnson (Eds.), Women's Studies: An Interdisciplinary Collection. Westport, Connecticut: Greenwood Press, 1978, 21-29.

Borod, J. "The Impact of a Women's Studies Course on Perceived Sex Differences, Real and Ideal Self-Perceptions, and Attitudes Towards Women's Rights and Roles. Doctoral dissertation, Case Western Reserve University, 1975.

Bose, C., Steiger, J., and Victorine, P. "Evaluation: Perspectives of Students and Graduates." Women's Studies Newsletter, 1977, 5, 6-7.

Bruner, J. The Process of Education. New York: Vintage Books, 1963.

Brush, L., Gold, A., and White, M. "The Paradox of Intention and Effect: A Women's Studies Course. Signs: Journal of Women in Culture and Society, 1978, 3 (4), 870-883.

Centra, J.A. "Evaluating College Teaching." National Conference on Higher Education Proceedings. Current Issues in Higher Education, 1972, 27, 225-233.

Centra, J.A. "Colleagues as Raters of Classroom Instruction." Journal of Higher Education, 1975, 46 (1), 327.

Coffman, S. "The Effect of Introductory Women's Studies Classes on Sex-role Stereotypes." Doctoral dissertation, University of Washington, 1978.

Coop, R.H., and Sigel, I.E. "Cognitive Style: Implications for Learning and Instruction." Psychology in the Schools, 1971, 8 (2), 152-61.

Costin, F., Greenough, W.T., and Menges, R.J. "Student Ratings of College Teaching: Reliability, Validity and Usefulness." Review of Educational Research, 1971, 41, 511-535.

Cross, K.P. "Not Can, But Will College Teaching Be Improved?" In J.A. Centra (Ed.), Renewing and Evaluating Teaching. San Francisco: Jossey-Bass, 1977.

Davis, B.H. "Finding New Forms: Student Autonomy in a Patriarchal University." University of Oklahoma, 1978.

Davis, P.K. and others. "Final Evaluation of GCI-336: Images of Women in Literature." University of Minnesota, General College Studies, 5 (3), 1975-76. (ERIC ED 124 240).

Davis, R.H. "Learning by Design." In J.A. Centra (Ed.), Renewing and Evaluating Teaching. San Francisco: Jossey-Bass, 1977.

Dewey, John. "My Pedagogic Creed." In R. Archambault (Ed.), John Dewey on Education: Selected Writings. Chicago: University of Chicago Press, 1969.

Doyle, Walter. Paradigms in Teacher Effectiveness Research. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, D.C., April 1975.

Dubin, R., and Tareggia, T.C. The Teaching-Learning Paradox: A Comparative Analysis of College Teaching Methods. Eugene: University of Oregon Center for the Advanced Study of Educational Administration, 1968.

Eileenchild, M. (Informal discussion.) Portland State University, Portland, Oregon, May 1979.

Elovson, A., and Cockcroft, I. "The Reported Impact of Women's Studies Courses on Students' Lives." Del Mar, California, 1977.

Farley, J. "Who Should Teach Women's Studies? An Answer from Course Evaluations." Teaching Sociology, 1978, 5 (4), 377-386.

Feldman, K., and Newcomb, T.M. The Impact of College on Students. San Francisco: Jossey-Bass, 1969.

Ferber, M.A., and Huber, J.A. "Sex of Student and Instructor: A Study of Student Bias." American Journal of Sociology, 1976, 80 (4), 949-963.

Fireman, J. "Reflections on Teaching Women's History: First Down and Goal to Go!" Journal of the West, 1973, 12, 197-211.

Flanders, N.A., and Simon, A. "Teacher Effectiveness." In R.L. Ebel (Ed.), Encyclopedia of Educational Research. New York: Macmillan, 1969.

Frey, P.W., Leonard, D.W., and Beatty, W.W. "Student Ratings of Instruction: Validation Research." American Educational Research Journal, 1975, 12, (4), 435-447.

Frieze, I.H., Parsons, J.E., Johnson, P.B., Ruble, D.N., and Zellerman, G.L. Women and Sex Roles: A Social Psychological Perspective. New York: Norton and Co., 1978.

Gage, N.L. Teacher Effectiveness and Teacher Education: The Search for a Scientific Basis. Palo Alto: Pacific Books, 1972.

Gould, M. "Teaching Sociology." Sociologists for Women in Society, Special Insert, 1978, 8 (2), 4.

Guttentag, M. "Subjectivity and Its Use in Evaluation Research." Evaluation Magazine, 1973, 1 (2), 60-65.

Guttentag, M., Brush, L., Gold, A., Mueller, M., Tobias, S., and White, M. "Evaluating Women's Studies: A Decision-Theoretic Approach." Signs: Journal of Women in Culture and Society, 1978, 3 (4), 884-890.

Harris, M. "Sex-role Stereotypes and Teacher Evaluations." Journal of Educational Psychology, 1975, 67 (6), 751-756.

Harris, M. "The Effects of Sex, Sex-stereotyped Descriptions, and Institution on Evaluation of Teachers." Sex Roles, 1976, 2 (1), 15-21.

Henderson, J.E., and Lanier, P.E. "What Teachers Need to Know and Teach." Journal of Teacher Education, 1973, 24 (1), 3-16.

Hersh, B. "On Teaching Women's Studies." Program Notes, Women's Studies Program, Northeastern Illinois University, 1979, 4 (1).

Hildebrand, M. "How to Recommend Promotion for a Mediocre Teacher Without Actually Lying." Journal of Higher Education, 1972, 43, 44-62.

Hoffman, N.J. "Reading Women's Poetry: The Meaning and Our Lives." College English, 1972, 34 (1), 48-56.

Howe, F. "Identity and Expression: A Writing Course for Women." Female Studies: No. 2, KNOW, Inc., 1970.

Howe, F. "Teaching Literature." Unpublished manuscript, 1976.

Howe, F. Seven Years Later: Women's Studies Programs in 1976. Washington, D.C.: National Advisory Council on Women's Educational Programs, 1977.

Kashak, E. "Sex Bias in Student Evaluations of College Professors." Psychology of Women Quarterly, 1978, 2 (3), 235-241.

Kritek, P., and Glass, L. "Nursing: A Feminist Perspective." Nursing Outlook, 26 (3), 182-6.

Kuhn, M.A., and McPartland, P.S. "An Empirical Investigation of Self-Attitudes." American Sociological Review, 19, February 1954, 68-76.

Kuhn, T.S. The Structure of Scientific Revolutions. Chicago: University of Chicago Press, 1970.

Lerner, Gerda. (Informal discussion.) Lewis and Clark College, Portland, Oregon, May 1979.

McKeachie, W.J., Lin, Y., and Mann, W. "Student Ratings of Teacher Effectiveness: Validity Studies." American Educational Research Journal, 1971, 8, 435-445.

Mangione, T.W., and Wiersma, G.E. "An Evaluation of the UMass/Boston Women's Studies Program." 1974.

Marsh, H.W., Fleiner, H., and Thomas C.S. "Validity and Usefulness of Student Evaluations of Instructional Quality." Journal of Educational Psychology, 1975, 67 (6), 833-836.

Martyna, W. "On Teaching Language and the Sexes." Women and Language News, Department of Linguistics, Stanford University, 1976.

Measurement Services Center. "Achievement Motivation in Women: Implementation and Evaluation of an Experimental Course." University of Minnesota, 1975.

Measurement Service Center. "Ego Development and Attitudes Toward Women in Women's Studies Students." University of Minnesota, 1976.

Morgan, E. "On Teaching Women's Studies." University of Michigan Papers in Women's Studies. Ann Arbor: Women's Studies Program, 1978, 27-34.

Register, C., Jones, J., Noble, B., and Faunce, P. "The Women's Studies Senior Seminar: 1976-77." Women's Studies Program, University of Minnesota, Minneapolis, 1978.

Rodin, M., and Rodin, B. "Student Evaluation of Teachers." Science, 1972, 177, 1164-1166.

Ruble, D., Croke, J., Frieze, I., and Parsons, J. "A Field Study of Sex-role Attitude Changes in College Women." Journal of Applied Social Psychology, 1975, 5, 110-117.

Schniedewind, N. "Feminist Values: Guidelines for Teaching Methodology in Women's Studies." 1978-79.

Schramm, S.S. "Women's Studies: Its Focus, Idea Power, and Promise." In K. Blumenhagen and W. Johnson (Eds.), Women's Studies: An Interdisciplinary Collection. Westport, Connecticut: Greenwood Press, 1978, 3-11.

Scott, C.S. "Some Remarks on 'Student Ratings: Validation.'" American Educational Research Journal, 1975, 12 (4), 444-447.

Scott, R., Richard, A., and Wade, M. "Women's Studies As Change Agents.'" Psychology of Women Quarterly, 1977, 1 (4), 377-379.

Shueman, S., and Sedlacek, W. "An Evaluation of a Women's Studies Program." Counseling Center Research Report #6-76, University of Maryland, College Park, 1976.

Signs: A Journal of Women in Culture and Society, 1975, 1 (1), v-viii.

Smith, B.O. (Ed.). Research in Teacher Education. Englewood Cliffs, N.J.: Prentice-Hall, 1971.

Socklaff, A.L., and Papacostas, A.C. "Uniformity of Faculty Attitude Towards Effective Teaching in Lecture/Discussion Courses." Journal of Educational Measurement, 1975, 12 (4), 281-293.

Stitzel, J. "Conve·sations." New Directions in Teaching, 1977, 5 (4), 14-17.

Stoper, E. "Problems and Opportunities in Teaching Women's Studies." California State University/ Hayward, undated.

Speizer, J. "An Evaluation of the Changes in Attitudes Toward Women Which Occur As a Result of Participation in a Women's Studies Course." Doctoral dissertation, Boston University School of Education, 1975.

Talburtt, M. and others. Project WELD Women's Education: Learning and Doing, a Final Report. Ann Arbor: Formative Evaluation Research Associates, Inc., 1977.

Thames, A.M. "Women's Studies in Three Institutions of Higher Education in California." Dissertation, University of California, Los Angeles, 1975.

Thorne, B., and Purrington, B. "Teaching a Course on Gender." Sociologists for Women in Society, Special Insert, 1978, 8 (2), 10-11.

Tobias, S., Brush, L., and Gold, A. "Wesleyan Conference Considers How to Evaluate Women's Studies." Women's Studies Newsletter, 1973, 1 (4), 11.

Wilson, J.J. "Women's Studies at the University of Puget Sound: Evaluation Report." 1976.

Winne, P.H., and Marx, R.W. "Reconceptualizing Research on Teaching." Journal of Educational Psychology, 1977, 69 (6), 668-678.

Wolf, J. "What Do You Teach in Women's Studies." Sociologists for Women in Society Newsletter, Special Insert, 1978, 8 (2), 12.

Women's Studies Program. "Self-survey Report." University of Minnesota, 1978.

Yamoor, C.M., Mischenko, M., and Martin, N.J. Life Styles of Educated American Women: Course Evaluation. Minneapolis: University of Minnesota Measurement Center, 1973.

Yates, G.G. "Women's Studies in Its Second Phase." Women's Studies Newsletter, 1977, 5 (1+2), 4-5.

APPENDIX

The following is a selection of responses to our survey of values in women's studies teaching.* Out of 285 programs contacted, 143 responded. Three individual teacher responses were counted maximally per institution. N = the number of individual women's studies teacher responses per question.

	N =	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Course content should relate directly to student lives	256	49%	41%	7%	3%	--
Course content should be concerned primarily with the subject matter	243	34%	54%	4%	7%	1%
Manner or style of teaching is as important as what is taught	256	51%	34%	9%	6%	--
Courses should contribute to personal growth of students	257	69%	28%	3%	--	--
Courses should contribute to acquisition of intellectual skills	258	78%	21%	1%	--	--
Courses should contribute to political development of students	256	38%	37%	23%	2%	--

Of the 50 percent (N=125) who indicated that their values in women's studies teaching had changed since they began (total is less than 100 percent):

- 30 percent (N=34) said they value student capacities, interaction, personal growth, and group process in the classroom more;

- 35 percent (N=39) said they value intellectual rigor, cognitive content, and skills of learning more.

*The survey of women's studies teachers was intended to aid in the preparation of this planning study. The authors make no claims to original or scientific research.

Of the 68 percent (N=169) who indicated agreement with the statement that feminist teaching styles or approaches exist (total is more than 100 percent):

- 54 percent (N=89) mentioned feminist perspectives on classroom process and organization;

- 22 percent (N=36) mentioned feminist approaches to subject matter;

- 20 percent (N=34) mentioned explicit connections between personal experience and reading materials;

- 16 percent (N=26) mentioned validating humanistic qualities in subject matter, pedagogical style, and how students are regarded;

- 14 percent (N=24) mentioned focus on women's experience, values, contributions, and relationship of course contents to feminist issues, including class and race;

- 12 percent (N=20) mentioned openness to dialogue, challenge, and new ideas;

- 21 percent (N=35) listed other attributes.

Some 94 percent (N=239) of the respondents indicated that their women's studies classes are regularly evaluated. When queried about level of satisfaction with current evaluation techniques, 60 percent (N=151) indicated satisfaction or strong satisfaction; 39 percent (N=98) indicated neutrality or dissatisfaction; and 1 percent (N=2) indicated strong dissatisfaction.

* U. S. GOVERNMENT PRINTING OFFICE : 1980 623-933/1544